

Homework 8

April 26, 2022

Jose Carlos Munoz

2-8

When the value of the entries are 0 or 1, we have two possible outcomes. The loss for each entry with a value of 1 is $L_i - \log(\phi((M)_{ij}))$. When the entry value is 0, the loss will be $L_i - \log(1 - \phi((M)_{ij}))$. Here $\phi()$ is the sigmoid function

When the Entries are a value between 0 and 1 we have a combination of the two losses at 0 and 1. This means that the entry that the function of the loss for this values will be $L_i = -y_i \log(\phi((M)_{ij})) - (1 - y_i) \log(1 - \phi((M)_{ij}))$. Where y_i is the entry value.

4-6

Since the training data and the testing data are close in accuracy, we can conclude that the model is underfitting. To fix this, it would be good if we add more units that have non linearity

4-10

The classification accuracy on the training data get worse as we increase the training data size. This is because the model can not remember the training data with limited capacity.

For the average loss, it increases as we have more data because the model is less specific for a particular training instance.

The testing data accuracy typically increases as we have more training data. So, at some point there will be a moment in which the training data and the testing data will be around the same accuracy.